ESTIMATION OF FUTURE WHEAT PRODUCTION FROM RAINFALL[1]

By Prof. HOMER J. HENNEY

[Kansas State College, Manhattan, Kans., June 1935]

This paper deals with methods of estimating Kansas winter wheat production in crop-reporting districts, several months before harvest, on the basis of previous rainfall.

The discussion is in four parts: (1) The general procedure or technic in determining the more influential periods of rainfall; (2) important factors in each crop reporting district; (3) some of the independent factors of lesser importance; and, (4) the possibilities of making an estimate 12 to 24 months prior to harvest.

PROCEDURE AND METHOD

These studies began with efforts to correlate the subsoil moisture data for land used in experiments at the Fort Hays branch of the Kansas Agricultural Experiment Station (1) with the yields of wheat secured in the county where the Hays station is located. The study was later expanded into correlation analysis of the relation between rainfall and wheat yields by crop-reporting districts in Kansas. The original purpose was to find some means of estimating by January, February, or March, the quantity of wheat to be harvested in Kansas in late June and July. Errors made in the spring months in judging future wheat prices have been attributed to month-to-month changes in estimates of the Kansas wheat crop. Correlation studies of prices and estimates of supplies show that the change in the estimates of Southwest winter wheat production in the spring months is the most important factor causing month-to-month price changes. It has been common knowledge for years that fall rainfall is a dominant factor in winter-wheat production; i. e., rainfall above normal has been associated with yields above normal in the hard winter wheat belt, but with yields below normal in the soft winter wheat belt, and vice versa. (3)

Since the cost of obtaining data on subsoil moisture for a depth of 6, or even 3, feet by counties in Kansas would be prohibitive, and since soil moisture is dependent upon rainfall, it seemed logical that rainfall data and county yields might show a correlation. At first, calculations were based upon rainfall at one station in each county and yields in that county. This was extended to include five counties in a group. The final correlations were based upon data from the nine crop reporting districts that are set up by the Government crop reporting service. Yield and production data are more often referred to by districts, and if the estimates based upon rainfall distribution are by districts it is easier to compare them with the estimates made by interpreting a condition figure issued either by the Government or by private agencies.

Some recent studies of rainfall and crop yields made by Production Credit Association areas indicate the State of Kansas might more logically be divided into type-of-farming areas (2). Soil types are one of the important factors in determining the type-of-farming areas, and the soil reaction to rainfall distribution varies greatly within the present crop-reporting districts.

For the purpose of this paper Y will refer to the dependent factor and X_1, X_2, etc., to the independent factors. In the studies herein reported, two general methods of making the estimate of the yield for the State were employed. The first method makes use of the independent factor in each district that shows the highest single relationship, as X_1, and this estimate is improved by applying the X_2 factor that gives the greatest reduction in the errors resulting from the first factor. The second independent factor, X_2, may not be the single factor that, used alone, gives the second best single relationship. No more than three independent factors together have been used. It might be possible to use more than three factors, especially if one started with an estimate 12 months prior to harvest and changed the estimate as combinations of monthly rainfall became available. The data prior to 1920 are not as complete and probably not as accurate as data for years since 1920, and this has necessitated the use of the shorter series of data which tends to limit all the more the number of constants one can use in the analysis.

The second method of estimating was to determine the simple relationships among the 3, 4, or 5 best factors. The modal estimate, if it could be called a modal with so few numbers, is used as the estimate for that district. These district estimates are totaled for the State average.

The first method is more scientific and should be more accurate over a period of years. It takes more time to keep that method up to date as new factors enter the situation, and it is of no more value for practical forecasting than the second method.

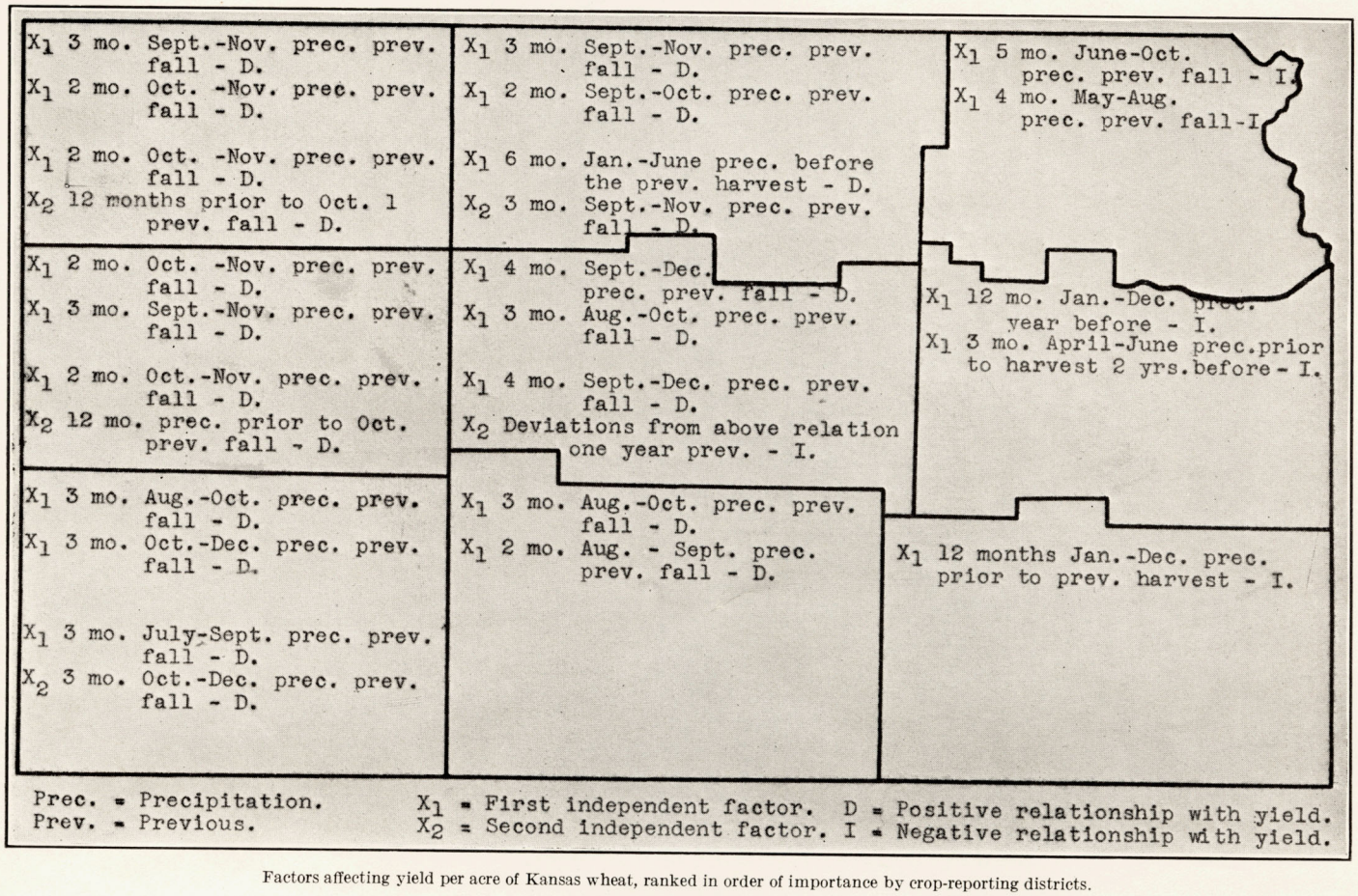
FACTORS AFFECTING YIELDS

The next phase of the study has to do with the more important factors that affect the yield in each district. The accompanying map of Kansas shows the crop-reporting districts (fig. 1). The State of Kansas is divided into thirds from west to east and from north to south, making nine districts. The northern three districts are numbered 1, 2, and 3 from west to east; the middle three, 4, 5, and 6, from west to east; and the southern three, 7, 8, and 9, from west to east.

Taking up the districts in numerical order, we find the most important single factor in district 1 is the rainfall during September, October, and November. The relation is direct with an index of correlation (4) of $+0.872$ and a standard error of estimate of ± 2.28 bushels per acre. The October-November or the September-October rainfall has almost as high a relation with yield. The earliest estimate for this area can be made about July 15,[2] before seeding the crop, by using the rainfall for 12 months prior to the previous harvest, which has a direct relation with yield per acre of $+0.616$. The direct relation of the rainfall so far ahead of the crop probably is due to the changes in the subsoil moisture during this period. In this district, both the acres planted and the acres harvested are related to rainfall, so that the direct relation between rainfall in the fall and total production is almost as high as the relation between rainfall and yield per acre. Whenever an area has this inner relationship, it is possible to arrive at an estimate several months earlier than when it is necessary to estimate the yield per acre

[1] Homer J. Henney, department of agricultural economics, Kansas State College of Agriculture and Applied Science, Manhattan, Kans.

[2] The 15th of the month following the month of latest rainfall is used instead of the 1st of the month following. Many times an estimate is possible by the 5th.



Factors affecting yield per acre of Kansas wheat, ranked in order of importance by crop-reporting districts.

first and then multiply by the acres sown after the data on acres sown are released in December.

In district 2 the highest relation of yield with a single factor is with the rainfall during the September-November period; i. e., near seeding time. The relation, by using the September-October or the September-December periods, is almost as high as with the September-November period. In this district, the rainfall for the 12 months prior to the beginning of the dominant [3] period, used as a second independent factor, did not improve the estimates. The earliest estimate for this area can be made on July 15 just after the harvest of the previous crop. The rainfall for both the 6-month period and the 12-month period, prior to the harvest of the previous crop, bears a direct relation to yield. The errors in estimating from the January to June rainfall during the growing season of the previous crop are reduced by including the September-November rainfall; but this estimate, of course, cannot be made until about December 15, which is not as early as an estimate can be made by using the September-October rainfall as a single factor. The index of correlation based upon the September-November rainfall was +0.825, and by the addition of the January-June rainfall as an X_2 factor it was +0.852. The rainfall for the calendar year before harvest gave a correlation of +0.80.

In district 3, an inverse relationship appears to exist between the rainfall and yield per acre, and also between rainfall and total production. The rain during the period June to October of the previous year seems to be the most important of any during the 12 months prior to harvest. Periods longer than 5 months seem to be more important than periods of less than 5 months. When any 12-month period is considered, the period 2 years prior to harvest seems to exercise a greater effect than the year immediately prior. In other words, the rain that fell for the 12 months prior to the harvest of the previous crop seems to be more important in determining yields than rainfall during the 12 months immediately prior to a current crop.

When periods of 1 to 2 years are considered, the rainfall tends to shift acreages and yields per acre in the same direction, or the relation between the rainfall and total production is higher than the relation between rainfall and yield per acre. Just why this is true is still unknown. Some agronomists have suggested that the heavy or above-normal rains from 1 or 2 years may have leached out some of the soil nutrients, and in years of below-normal rainfalls more plant food is left for the plant to absorb. So long as rainfall is sufficient to permit the plant to absorb the remaining soluble food, the plant will react to the above-normal supply of food retained in the soil in the drier years. In the eastern third of Kansas, and probably in the soft winter wheat belt, the rainfall in below-normal periods is still sufficient to permit some food assimilation by the plant. In the western districts of Kansas, the below-normal rainfall periods do not always permit the plant to absorb the above-normal supplies of nutrients left in the soil. For that reason there tends to be a direct relation between rainfall for the previous 1 to 2 years and production. The errors in estimating are greater in district 3 and in district 9 than in most other areas. The best single relationships were with rainfall at seeding time 1 year previous and 2 years previous to harvest.

[3] Dominant period refers to the one period that by simple correlation shows the highest relationship with yield.

In district 4 the correlation between rainfall and yield per acre is distinctly direct and the correlation between acre-yield and rainfall before the previous fall tends to be direct, thus indicating the influence of the subsoil moisture. The October–November precipitation tends to give the best relation with yield per acre, giving an index of correlation of +0.766. The errors of estimate can be reduced by using rainfall for 6- to 18-month periods prior to the October–November period, but the 12-month period just previous to the dominant period gives the greatest reduction in the errors. The earliest estimate for this district can be made on July 15 just after the previous harvest, by using the rainfall for the preceding January to June.

In district 5 the central district of the State, the best single factors to use, in order of importance, are the 4 months of August–November, the 3 months August–October, and the 4 months September–December of the previous fall, The index of correlation, with the August–November rainfall and the rainfall for the 12 months prior to the August period, is +0.717. The earliest estimate for this area can be made on September 15, by using the June–August rainfall; but this estimate needs considerable revision by adjusting with the August-November rainfall, which cannot be done before December 15. The yield per acre the previous year tends to show a negative relation with the current yield. In this district a yield below normal or lower than the previous year tends to give a yield above normal or above the previous year. A preliminary estimate on this basis can be made in June. Since 1921, with the previous September–December rainfall as X_1 and the yield per acre the previous year [4] as X_2, the index of correlation is +0.815. In this district more than in any other, the cumulative rainfall for 6 months seldom constitutes as good a factor as the rainfalls during 2 periods of 3 months each. Furthermore, the center month of a 5-month period, such as September for the July–November period, increases the correlation if coupled with the 2 preceding and 2 succeeding months into 2 factors, i. e., if September, e. g., be used in a 3-month total with July and August and again in a 3-month total with October and November. This may not be orthodox statistical procedure, and it is the author's hope that someone will be able to explain the reason for this higher relationship.

In this district, the rainfall for 12 months prior to the period that shows the greatest relation exhibits an inverse relation to yield, which is contrary to findings in some other districts, especially in the western third of the State. This fact tends to verify the theory that a yield that is above normal in bushels and above normal relative to the previous rainfall, depletes the soil of certain nutrients, which in turn causes a smaller yield the following year then the rainfall would indicate. It also lends support to the theory that abnormally low rainfall, especially if accompanied by abnormally high temperatures, may result in partial sterilization of the soil, which would in turn favorably influence productivity in the next season.

In district 6, the east central district, all the relationships worked out between rainfall and yield are inverse, and tend to be higher as the length of the rainfall period is increased. One of the best relationships found in this district is between yield and the rainfall during April, May, and June 2 years prior to harvest, i. e., the rainfall for the 3 months prior to July 15 can be used in estimating

[4] This is correlated not as yield the previous year, but as yield the previous year in bushels above or below the amount estimated on the basis of rainfall for that year.

the production 24 months later. The reason for this is not known, but is apparently due to some building up of surplus soluble plant food in dry years. Likewise, this is the only eastern area where the April to June rainfall just before harvest has any measurable relation to the yield.

In district 7 the relation of all rainfall periods to yield is positive, whether to a long period of 24 months or a short period of 2 months, and also whether the period chosen is 2 or 3 years or 2 months prior to harvest. This is true of no other area. Likewise, the October–December and the January–March rainfall periods show some measurable influence which, as a rule, is not detectable in most areas. The August–October 3-month period shows the highest relation, +0.536. This estimate is improved by using the rainfall for 12 months prior to the previous August. These two factors together give +0.817. A fairly good estimate in this area can be given on September 15, 9 months before harvest, by employing the July–September rainfall. Then on January 15 this can be improved by using the October–December rainfall as an X_2 factor, and some further reductions in the error can be made on June 15 by using as an X_3 factor the March–May rainfall. This district, like district 5, shows a tendency for the errors of estimate to be plus 1 year and minus the following year.

In district 8 the same type of relationships are found as in district 5, but the correlations are not as high. The fall rainfall has a direct relation to yield, but rainfall prior to the previous harvest tends to show an inverse relation. Combinations of the fall rainfall into two or more factors give the best single or multiple relationships. These estimates can usually be improved by using the rainfall for either the 6-month or 12-month period before the previous crop. The best single factor is the August–October rainfall which has a relation of +0.694 with yield per acre. The earliest estimate in this area is in July, when one can use the January–June previous rainfall which has an inverse relation to yield, the same as in district 5.

In district 9 two conflicting relationships are present. Apparently in a series of dry years the previous fall rainfall has a direct relation, and in a series of wet years the relation is inverse. This cannot be shown statistically as yet. Errors of estimate in this area are greater than in any other district and predictions appear to be of little value. The best single relation is with the rainfall for 12 months prior to the previous harvest.

OTHER FACTORS AFFECTING YIELDS

Research studies have been conducted in an attempt to reduce the errors in judging the yield of grain. The influence of the subsoil moisture has been checked by using the accumulation of rainfall for 1 to 3 years before the dominant period. As was shown in districts 3, 5, 6, and 8 this relation is inverse to the current yield; and in districts 1, 4, and 7 it is direct.

The theory that the spring rainfall must be forecast before one can improve the estimates has been checked. To date all that can be said is that with the exception of district 7 spring rainfall tends to increase the errors as often as it reduces them. On the other hand, it is common knowledge that spring rains are essential to the crop. In most districts, however, there is some direct relation between fall and spring rainfall. This inner relationship of the two periods might explain the lack of the need for using the spring rainfall. Further studies will no doubt make it possible to improve the estimates by using spring rainfall.

The concept that a large crop is followed by a small one has been tested for all districts, and such a relation was found in districts 5, 7, and 8 to some extent. Apparently, this is important where wheat occupies the larger percentage of the cultivated area. There appear to be possibilities in such studies if they could be made by type-of-farming areas.

The relation of temperature to yield has been checked to some extent, and apparently accounts for some of the greatest errors in years of below-normal rainfall. A below-normal rainfall, with an above-normal temperature in November and December, tends to give the same yield as above-normal rainfall and normal temperature. This accounts in a large measure for the 1914 wheat crop, which was about twice as large as the fall rainfall and acreage would indicate. The standard error of estimate of the relationships given vary from 2.5 to 3.5 bushels per acre in the acres sown. These errors are reduced but little when more than two factors have been used.

LONGER TIME ESTIMATES

If possible, it is desirable to estimate the supply far enough ahead for it to be of some value in adjusting winter wheat production. From the discussion already given it is seen that a preliminary estimate can be made in most areas 1 or 2 months before seeding; and a much better estimate can be made by the time the crop goes into the dormant period, i. e., November or December. Estimates before seeding must for the most part be based on cumulative rainfall for a 12-month period prior to the previous harvest. In the eastern third of the State, the estimates are an indication of the direction of variation from the average yield, and are of a practical value when one considers the advantages in having some idea so far in advance of harvest. Since 1920 there have been only 3 years of large errors in district 3, and only 2 years of large errors in district 6. The error in the other years was negligible for long-time production planning. Since a high percent of the United States winter wheat crop is produced under conditions similar to those in the different sections of Kansas, there appears to be a possibility of using weather records to indicate desirable adjustments in wheat production.

BIBLIOGRAPHY

(1) **Halstead, A. L., and Coles, E. H.**
A preliminary Report of the Relation Between Yield of Winter Wheat and Moisture in the Soil at Seeding Time. Jour. of Agr. Research 41:467–477, September 1930.

(2) **Hodges, J. A.**
The Effect of Rainfall and Temperature on Corn Yields in Kansas. Jour. of Farm Econ. 13:305–318, April 1931.

(3) **Alsberg, C. L., and Griffing, E. P.**
Forecasting Wheat Yields from the Weather. Wheat Studies, Food Research Institute, vol. V. no. 1, 1928.

(4) **Ezekiel, Mordecai.**
Methods of Correlation Analysis. John Wiley and Sons, New York, 1930.